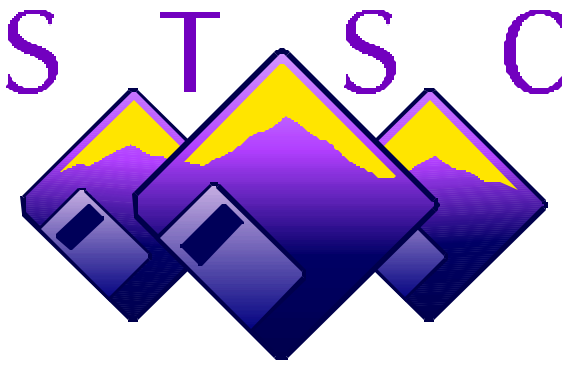


Software-Oriented Test and Evaluation (SOT&E)

National Defense Industrial Association (NDIA)

by Gregory T. Daich

August 15, 2001

Software Technology Support Center (STSC)

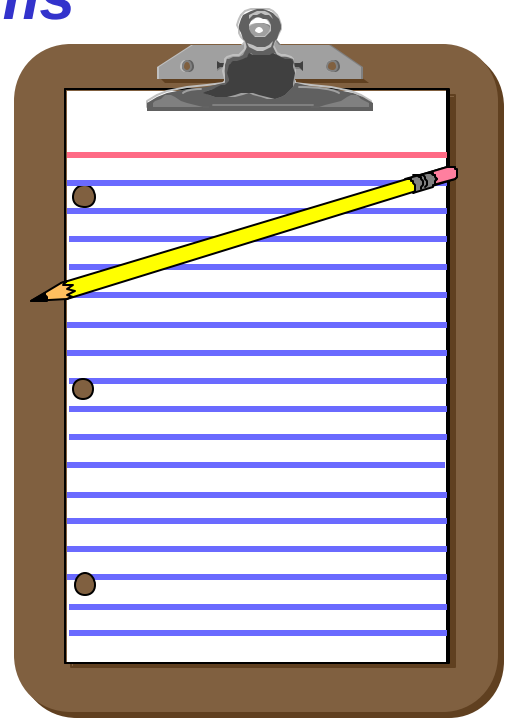
Science Applications International Corporation (SAIC)

Agenda

- Attendee Input
- Test and Evaluation (T&E)-Related Definitions
- The New 5000 Model
- T&E Supports Defense Acquisition
- Notional Partitioning of T&E Technologies
- Software T&E-Specific Technologies
- Fundamental Software Testing Technologies
- Acronyms
- References

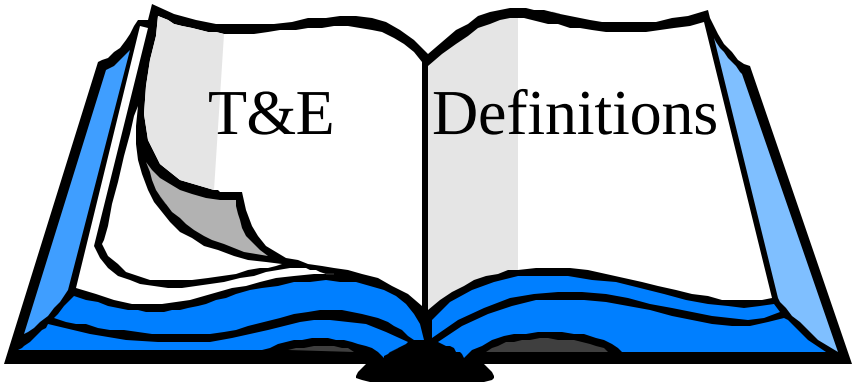
Attendee Input

- Identify *1 - 2 best practices* related to *software and T&E*
- Identify 1 - 2 *concerns* regarding *T&E of software-intensive systems*
- Please *turn ideas* in at end of presentation



Test & Evaluation-Related Definitions

- Test and Evaluation
- Developmental Test and Evaluation
- Operational Test and Evaluation
- Software-Oriented Test and Evaluation
- Verification, Validation
- Operational Effectiveness & Suitability



Test and Evaluation

- **Test**
  - "An activity in which a ***system or component is executed*** under specified conditions, the results are observed or recorded, and an ***evaluation is made*** of some aspect of the system or component." [IEEE610.12]
  - "A program, procedure, or process ***to obtain, verify, or provide data*** for determining the degree to which a system (or subsystem) meets, exceeds, or fails to meet its stated objectives." [TST101]

Test and Evaluation

- **Test**
  - "The term "test" denotes any project or program designed to ***obtain, verify, and provide data*** to ***evaluate, research, and develop*** (other than laboratory experiments); progress in accomplishing development objectives; performance and operational capability of systems, subsystems, and components; and equipment items" [AFOMan99-102]

Test and Evaluation

- **Evaluation**
  - Not in IEEE 610.12
  - "A systematic ***determination of the extent*** to which an entity ***meets its specified criteria***." [IEEE12207]
  - "The ***review, analysis, and assessment*** of data obtained from testing or other sources." [TST101]

STEPS

Test and Evaluation

- **Evaluation**
  - "The term "evaluation" denotes the ***review and analysis of data*** produced ***during*** current or previous ***testing*** and ***data obtained from test*** conducted by other government agencies and contractors, ***from operation and commercial experience***, or combinations thereof." [AFOMan99-102]

Test and Evaluation

- **Test and Evaluation**
  - "***Process*** by which a system or components ***provide information regarding risk*** and risk mitigation and empirical data to validate models and simulations. T&E permit, as assessment of the attainment of technical performance, specifications and system maturity ***to determine whether systems are operationally effective, suitable and survivable*** for intended use. There are two types of T&E -- Developmental (DT&E) and Operational (OT&E)." [DSMCGlossary]

Developmental Test and Evaluation

- "Developmental Test and Evaluation (DT&E), including contractor testing, is ***conducted to evaluate*** design approaches, ***validate analytical models, quantify contract technical performance and manufacturing quality***, ***measure progress in system engineering design and development***, ***minimize design risks***, and ***predict integrated system operational performance*** (effectiveness and suitability) in the intended environment and ***identify system problems*** (or deficiencies) to allow for early and timely resolution or correction." [AFPD99-1]

Developmental Test and Evaluation

- "Any engineering-type test used to ***verify status of technical progress***, ***verify*** that ***design risks are minimized, substantiate*** achievement of ***contract technical performance***, and ***certify readiness*** for initial operational testing. Development tests generally require instrumentation and measurements and are accomplished by engineers, technicians, or soldier operator-maintainer test personnel in a controlled environment to facilitate failure analysis." [DSMCGlossary]

Operational Test and Evaluation

- "The ***field test***, under realistic combat conditions, of any item of (or key component of) ***weapons***, ***equipment***, or ***munitions*** for the purpose of ***determining the effectiveness and suitability*** of the weapons, equipment or munitions for use in combat by typical military users, and the ***evaluation of the results of such test***."
  [10 USC Sec139]

Software-Oriented Test and Evaluation

- SOT&E term derived from [PN99-502]
- No “official” definition exists
- Both Air Force and Army use the term:
  “Software Test and Evaluation”
- My definition of SOT&E (augmented from IEEE 12207.0 definition of “evaluation”) which can serve well as an unofficial definition:
  - *A systematic determination of the extent to which a software entity meets its specified criteria through analysis, review, and execution of software products and processes.*

Software-Oriented Test and Evaluation

- **Air Force Operational Test and Evaluation Center (AFOTEC) Pamphlet 99-102 volumes are a series of software operational test and evaluation guides:**
  - **AFOTEC Pamphlet 99-102, Volume 2 -- Software Support Life Cycle Process Evaluation Guide** [AFOPam99-102v2]
  - **AFOTEC Pamphlet 99-102, Volume 4 -- Software Usability Evaluation Guide** [AFOPam99-102v4]
  - **AFOTEC Pamphlet 99-102, Volume 5 -- Software Support Resources Evaluation Guide** [AFOPam99-102v5]
  - **AFOTEC Pamphlet 99-102, Volume 6 -- Software Maturity Assessment Guide** [AFOPam99-102v6]
  - **AFOTEC Pamphlet 99-102, Volume 8 -- Software Operational Assessment Guide** [AFOPam99-102v8]

Software-Oriented Test and Evaluation

- **Army Pamphlet 73-7 contains Software Test and Evaluation Guidelines** [ARPam73-7]

Verification

- "***Confirmation*** by examination and provision of objective evidence ***that specified requirements have been fulfilled***.
- Notes
  1. In design and development, verification concerns the process of examining the result of a given activity to determine conformity with the stated requirements for that activity.
  2. "Verified" is used to designate the corresponding status."

  [IEEE12207]
- ***Verification assures "You built it right."***

  [CMMIC], p. 416

Validation

- "***Confirmation*** by examination and provision of objective evidence that the ***particular requirements for a specific intended use are fulfilled***. Notes:
  1. In design and development, validation concerns the process of examining a product to determine conformity with user needs.
  2. Validation is normally performed on the final product under defined operating conditions. It may be necessary in earlier stages.
  3. "Validated" is used to designate the corresponding status.
  4. Multiple validations may be carried out if there are different intended uses." [IEEE12207]
- ***Validation assures "You built the right thing?"*** [CMMIC], p. 416

Operational Effectiveness

"The overall ***degree of mission accomplishment*** of a system when ***used by representative personnel*** in the environment planned or expected (e.g., natural, electronic, threat, etc.) for ***operational employment of the system*** considering ***organization, doctrine, tactics, survivability, vulnerability, and threat (including countermeasures, initial nuclear weapons effects, nuclear, biological, and chemical contamination (NBCC) threats***)."

[DAD], Glossary

Operational Suitability

"The ***degree*** to which a system can be ***placed satisfactorily in field use*** with consideration being given to ***availability, compatibility, transportability, interoperability, reliability, wartime usage rates, maintainability, safety, human factors, manpower supportability, logistic supportability, natural environmental effects and impacts, documentation, and training requirements***."

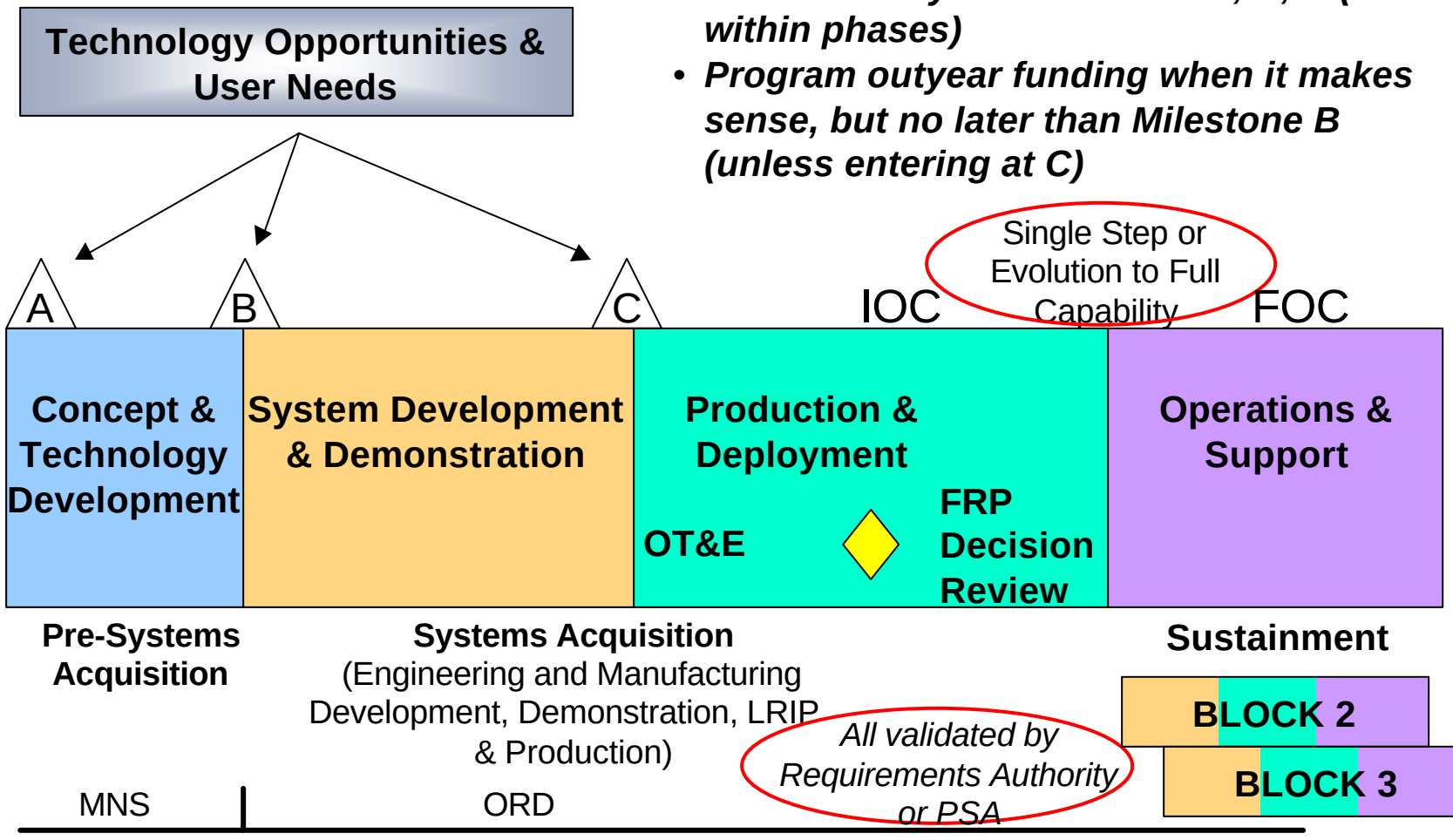
[DAD], Glossary

The New 5000 Model

- **Acronyms**
  - **Full Operational Capability (FOC)**
  - **Full Rate Production (FRP)**
  - **Initial Operational Capability (IOC)**
  - **Live Fire T&E (LFT&E)**
  - **Low Rate Initial Production (LRIP)**
  - **Mission Needs Statement (MNS)**
  - **Operational Requirements Document (ORD)**
  - **Principle Staff Assistant (PSA)**

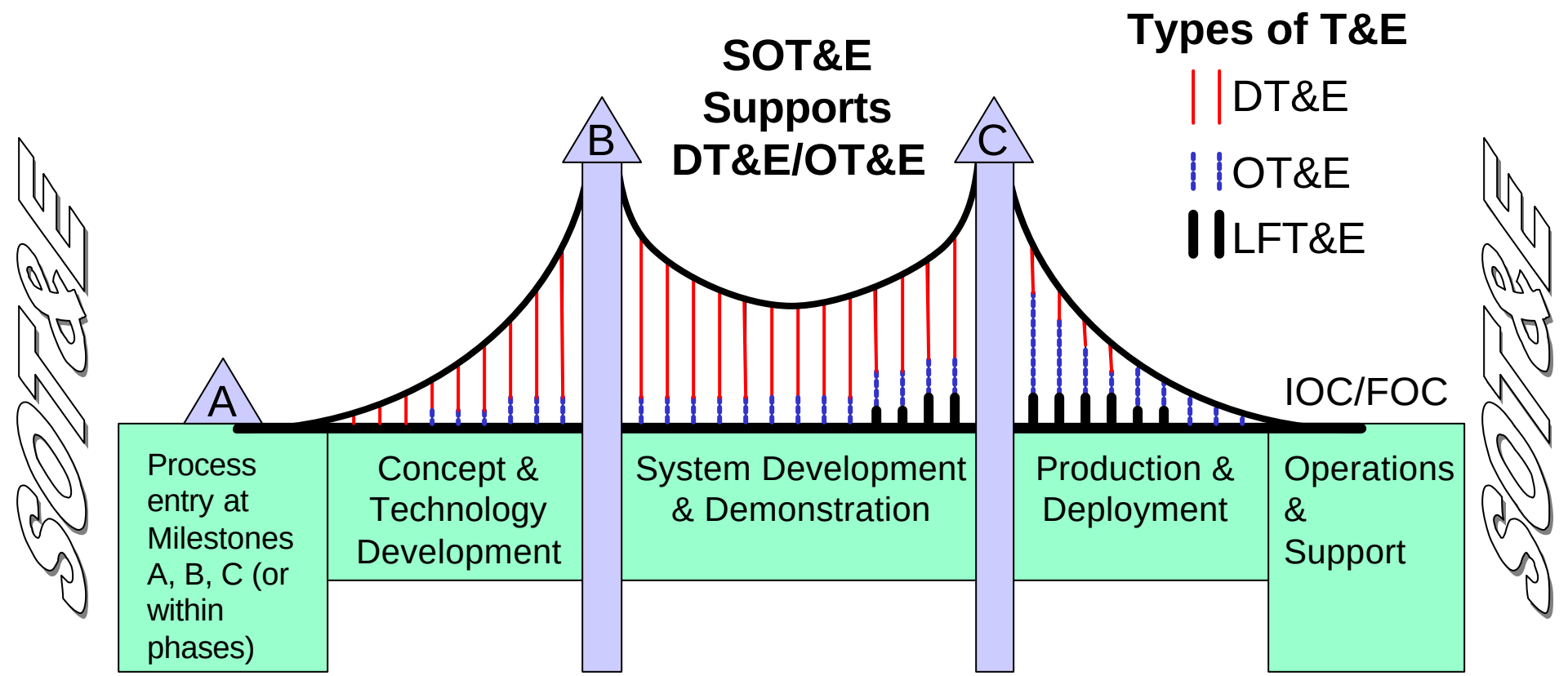
The New 5000 Model

- ***Process entry at Milestones A, B, C (or within phases)***
- ***Program outyear funding when it makes sense, but no later than Milestone B (unless entering at C)***



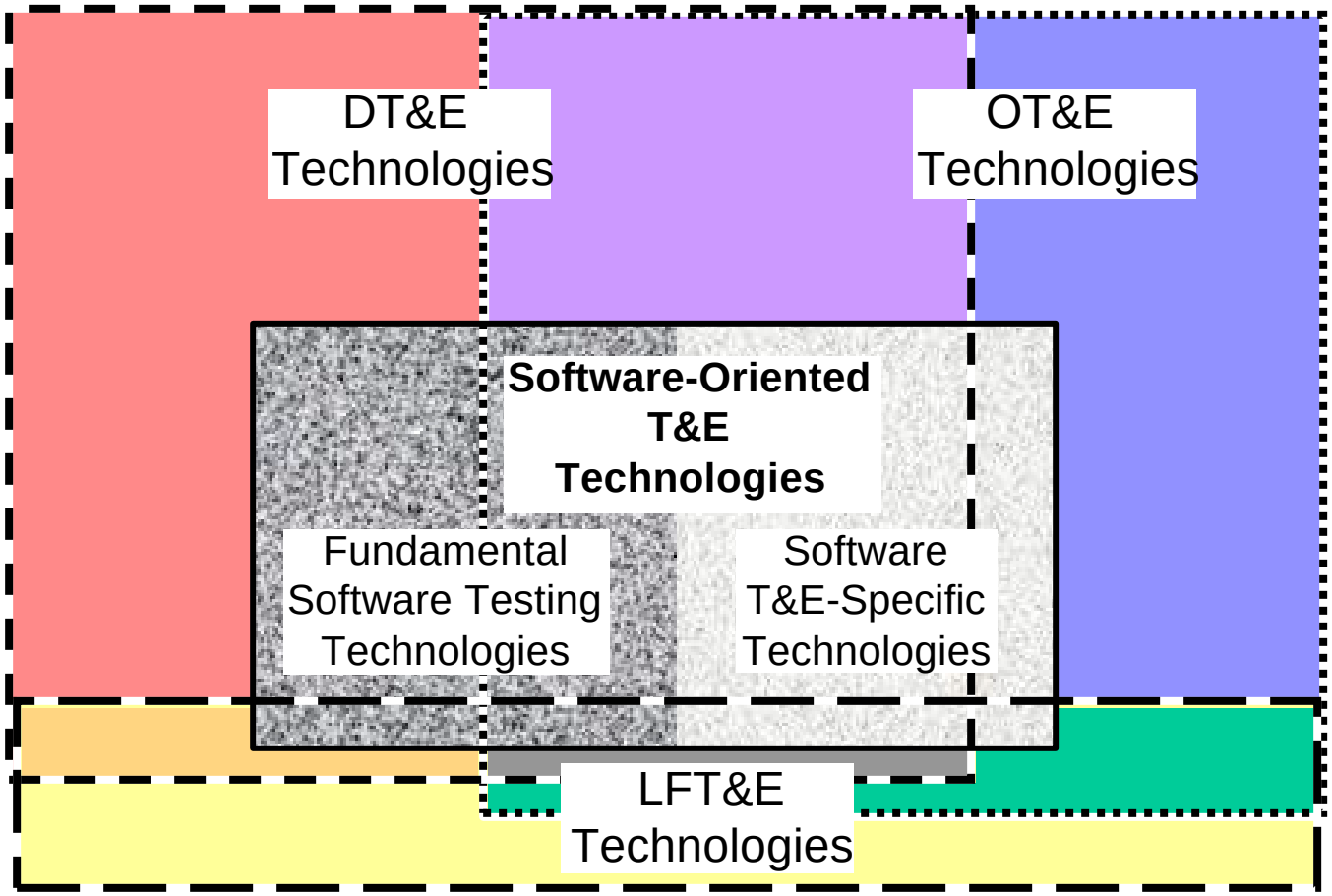
Relationship to Requirements Process

T&E Supports Defense Acquisition



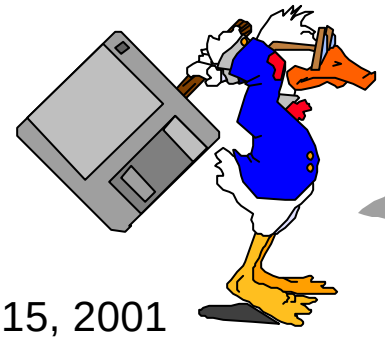
[TST202], Role of T&E, p2
[DoDI5000.2], Fig. F1

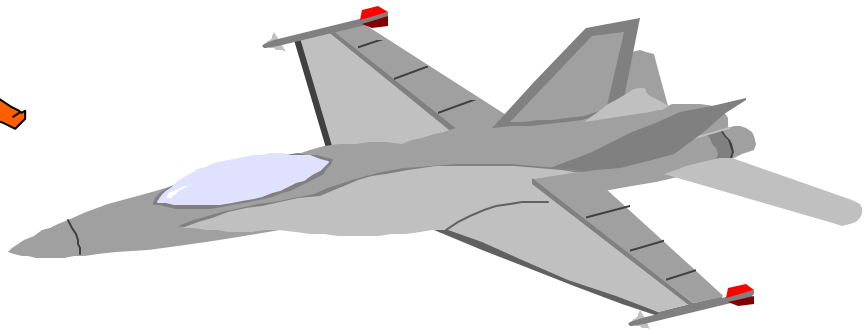
Notional Partitioning of T&E Technologies



Software T&E-Specific Technologies

- Key T&E Source Documents
- Test & Evaluation Master Plan (TEMP)
- Software-Specific T&E Levels
- Mission & Operational Requirements Coverage
- Acquirer-Developer Joint Reviews
- Testing Software-Intensive Commercial-off-the-Shelf (COTS) / Non-developmental Items (NDI)
- Software T&E-Specific Best Practices





Key T&E Source Documents

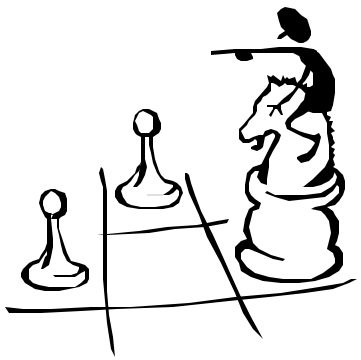
- **Title 10, U.S. Code, Section 139** [10 USC Sec139]
- **Department of Defense (DoD) Directive 5000.1** [DoDD5000.1]
- **DoD Regulation 5000.2-R** [DoD5000.2-R]
- **DoD Instruction 5000.2** [DODI5000.2]
- **Air Force (AF) Policy Directive 99-1** [AFPD99-1]
- **AF Instruction 99-101, 99-102** [AFI99-101],
- **AF Instruction 99-102** [AFI99-102]

Key T&E Source Documents

- **Army Regulation** [AR73-1]
- **Secretary of the Navy Instruction (SECNAVINST) 5000.2B** [SECNAVI5000.2B]
- **Marines Corp Order (MCO) 3960.2B** [MCO3960.2B]
- **IEEE/EIA 12207.0 Standard for Information Technology - Software Life Cycle Processes** [IEEE12207]
- **EIA/IEEE J-STD-016, Software Development Acquirer-Supplier Agreement** [J-STD-016]

Test & Evaluation Master Plan

- **Your program may not require a TEMP, but you will still likely need a Master Test Plan**
- **TEMP defines T&E strategy for DT&E, OT&E, LFT&E**
- **It shall relate program schedule, test management strategy and structure, and required resources to threshold and objective evaluation criteria derived from the ORD for**
  - **Critical Operational Indicators (COI)**
  - **Critical Technical Parameters (CTP)**
  - **Key Performance Parameters (KPP)**
  - **Operational Performance Parameters (OPP)**
  - **Major decision points**



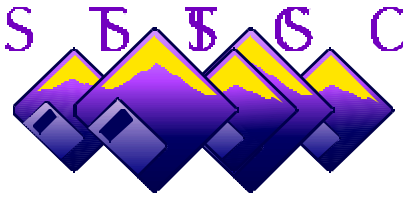
Test & Evaluation Master Plan

- **It shall include at least one CTP and one operational effectiveness issue for the evaluation of interoperability.**

  [DoD5000.2-R]
- **Consider using the Automated Test Planning System**
  - **Expert system that aids in preparing the TEMP**
  - **Access to latest Interim DoD 5000 Series documentation**
- **Each project establishes Integrated Product Team (IPT) to enable integrated process and product development**
- **Each service establishes an Operational Test Agency (OTA) with experience to support OT&E**

Atps.lnk

Software-Specific T&E Levels



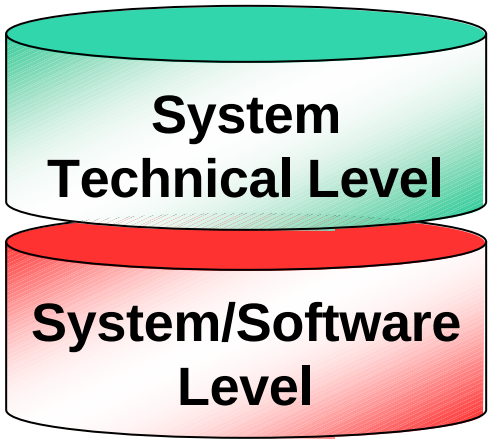
- **Acronyms**
  - **Concept of Operations (CONOPS)**
  - **Hardware Configuration Item (HWCI)**
  - **Need Determination Record (NDR)**
  - **Software Architectural Description (SAD)**
  - **Software Configuration Item (SWCI)**
  - **Software Design Description (SDD)**
  - **Software Product Description (SPD)**
  - **Software Requirements Description (SRD)**
  - **System Requirements Specification (SRS)**

    [IEEE12207]

Software-Specific T&E Levels

NDR (MNS)
CONOPS
(ORD, SRS)

OT&E Plan



SRS
SRD

DT&E Plan
(government)

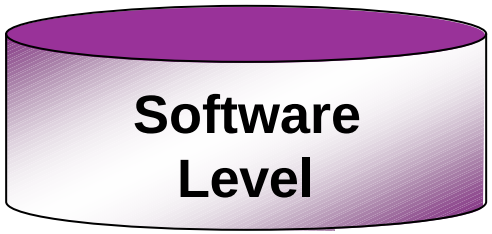
Test Plan (system
qualification test)

SWCI/HWCI Integration
Test Plan

SAD

SDD

SPD

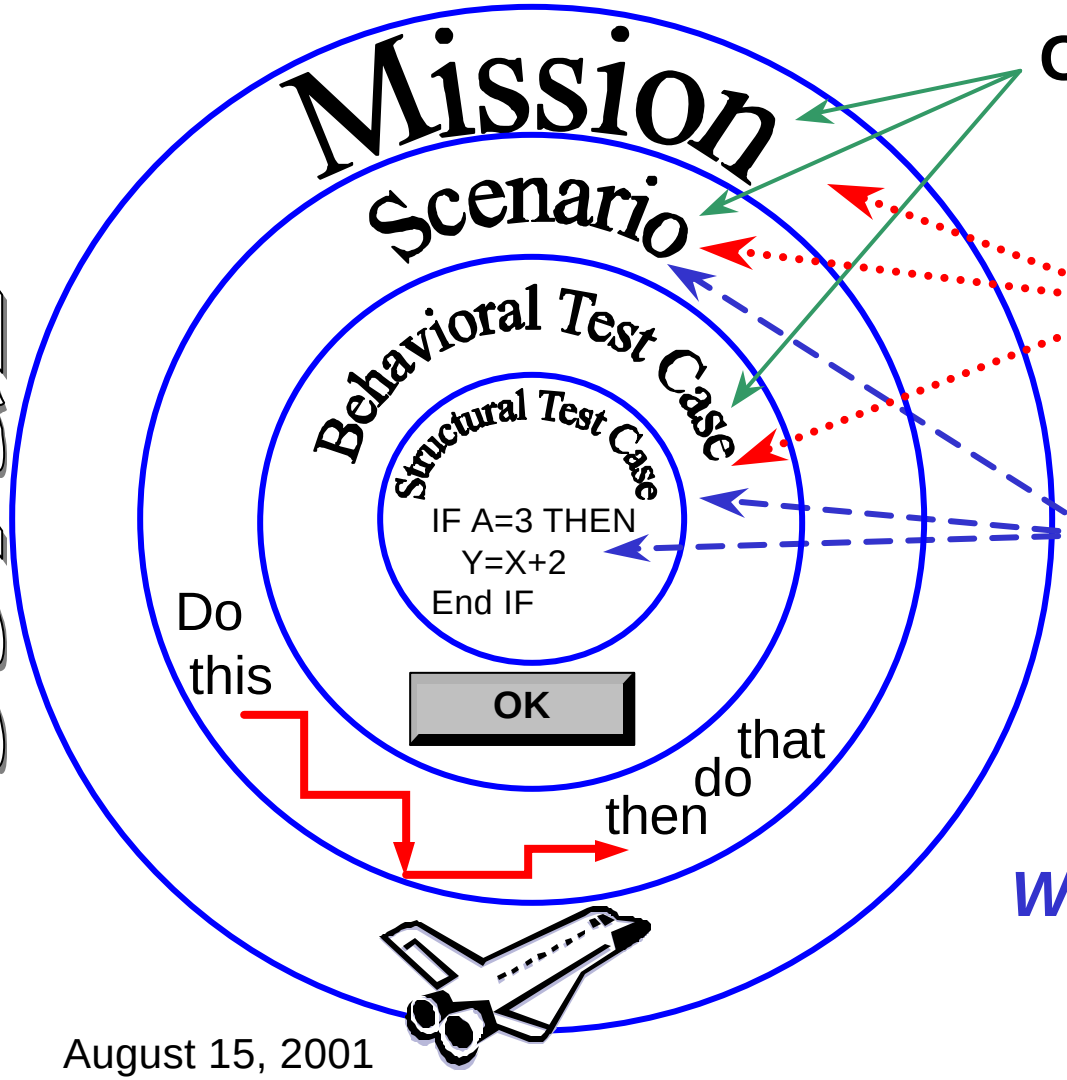


SWCI Qualification Test
Plan

Integration & Test Plan

Unit Test Plan

Software-Specific T&E Levels



OT&E/EOA/
IOT&E/FOT&E
(Acquirer/User)

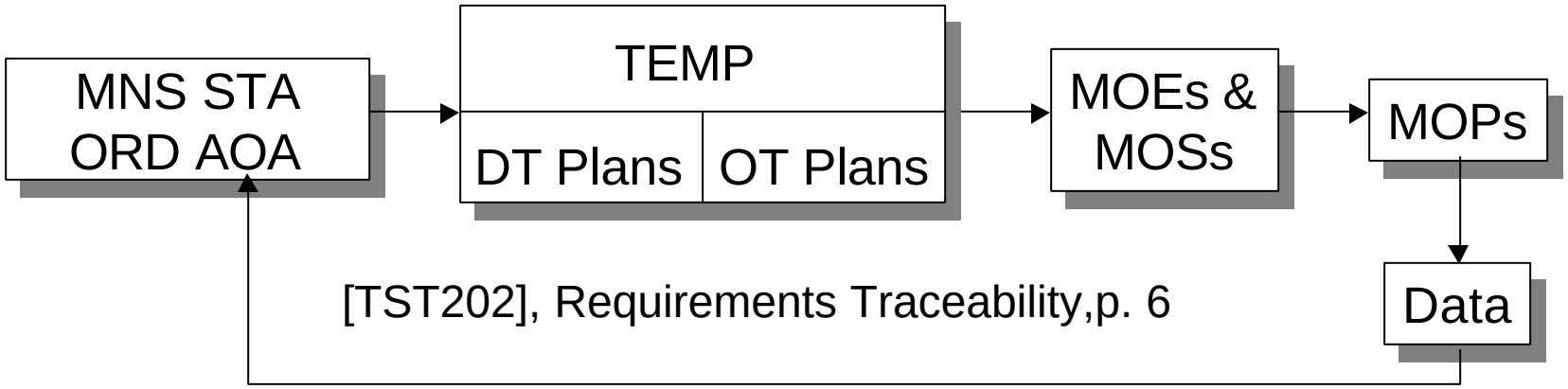
DT&E (government)
(Acquirer/User)

DT&E (contractor)
(Acquirer/Supplier/
Developer/User)

What is your T&E focus?

Mission & Operational Requirements Coverage

- **Requirements Correlation Matrix**
  - **Identifies Measures of Effectiveness (MOE), Suitability (MOS), and Performance (MOP) thresholds / objectives for each operational requirement derived from MNS, System Threat Assessment (STA), ORD, and / or Analysis of Alternatives (AOA)**
  - **E.g., Anthropometrics (body size) Accommodation**
    - Threshold: 34 - 40 in.
    - Objective: 32.4 - 41.4 in.

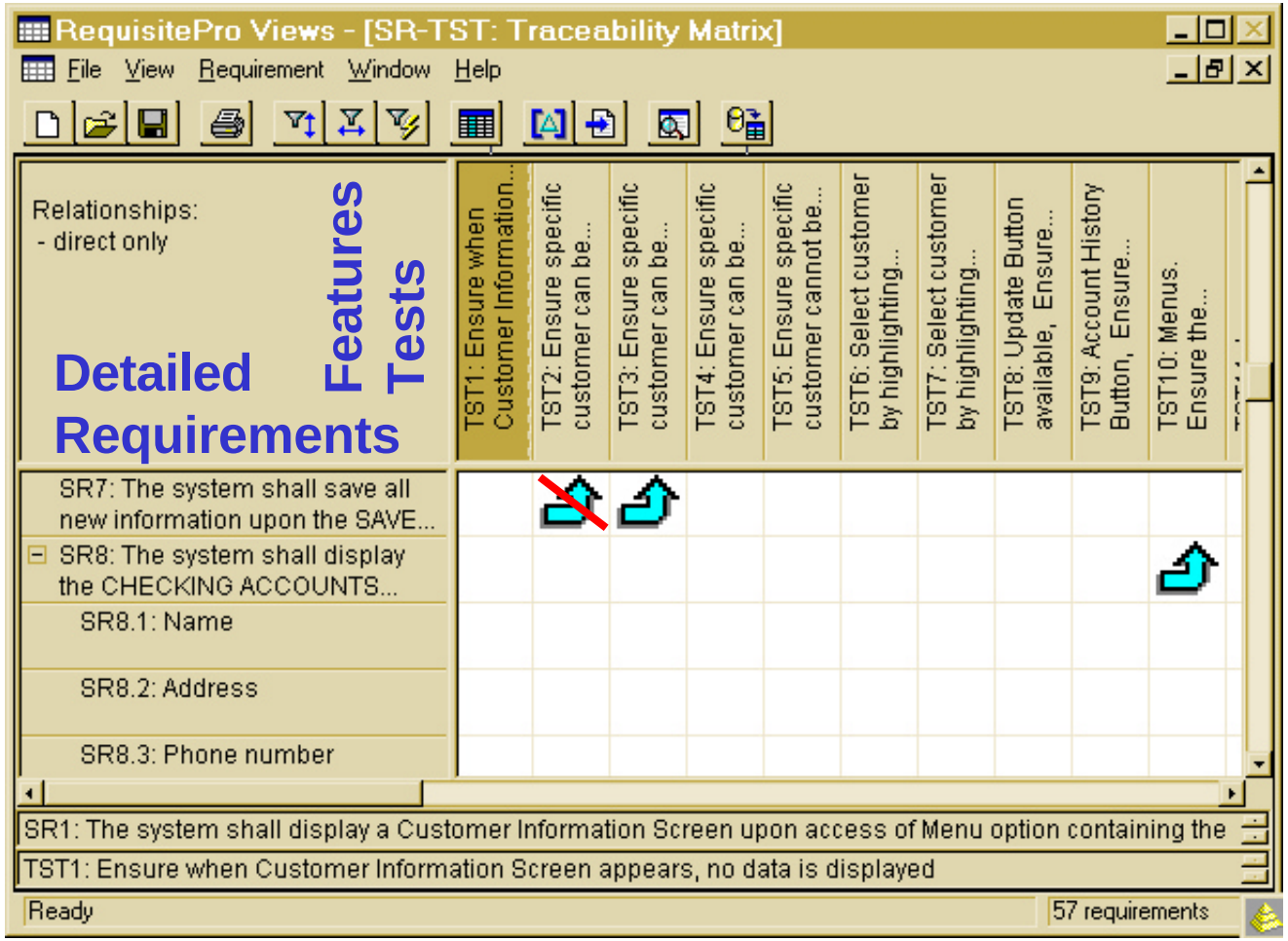


Mission & Operational Requirements Coverage

- **Sometimes traceability matrices trace high-level mission-oriented operational requirements to detailed technical requirements, to design components, and to code components.**
- **Some Requirements Verification Traceability Matrix (RVTM) or Verification Requirements Traceability Matrix (VRTM) map requirements to the type of evaluation methodology to be conducted, such as** ([TST101] p. 26):
  - Inspection, Demonstration, Analysis, Test
- **Effective software testing traces requirements to test cases**

Mission & Operational Requirements Coverage

Example Detailed Requirements to Feature Tests Requirements Tracing



Acquirer-Developer Joint Reviews

- **Acquirers should participate in developer's formal document reviews as much as possible**
  - **Prior to joint technical reviews such as Preliminary Design Reviews (PDR) or Test Readiness Reviews (TRR)**
- **Acquirers should audit developer document review database**
  - **To determine if disciplined (effective and efficient) document reviews are being conducted**
- **Acquirers should conduct disciplined document reviews on samples of deliverables**
  - **Especially if quality of developer document reviews is suspect**
  - **And to determine the level of document quality delivered**

Testing Software-Intensive COTS/NDI

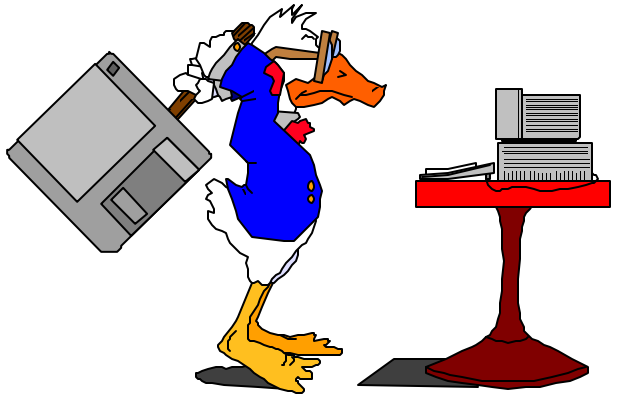
- **Have critical success factors and acceptance criteria been identified and are they agreed to by key players?**
- **Have appropriate changes been made to business processes?**
- **Have testers been involved early in evaluation and selection and in preparing a test plan?**
- **Has the scope of entire conversion and preparation effort been adequately planned?**
- **Has adequate training been planned early?**
- **Does vendor (supplier) have a current list of defects and does vendor provide test plans?**

Software T&E-Specific Best Practices

- Effective acquirer, developer, and joint project document review practices are used for
  - ORD, TEMP, Request for Proposal (RFP), Software Development Plan (SDP), SRS, System Test Plan (STP)
- Mission and operational requirements are traced to test cases
- Early developmental and operational test planning
- Early resource planning in TEMP
- Update ORD and TEMP regularly
- Understand and apply applicable DoD and industry directives, regulations, guidelines, & standards
- ...

Fundamental Software Testing Technologies

- Test Planning
- Test Case Design
- Test Execution
- Test Management
- Test Process Definition
- Test-Related Document Reviews
- Test Automation
- Fundamental Software Best Practices



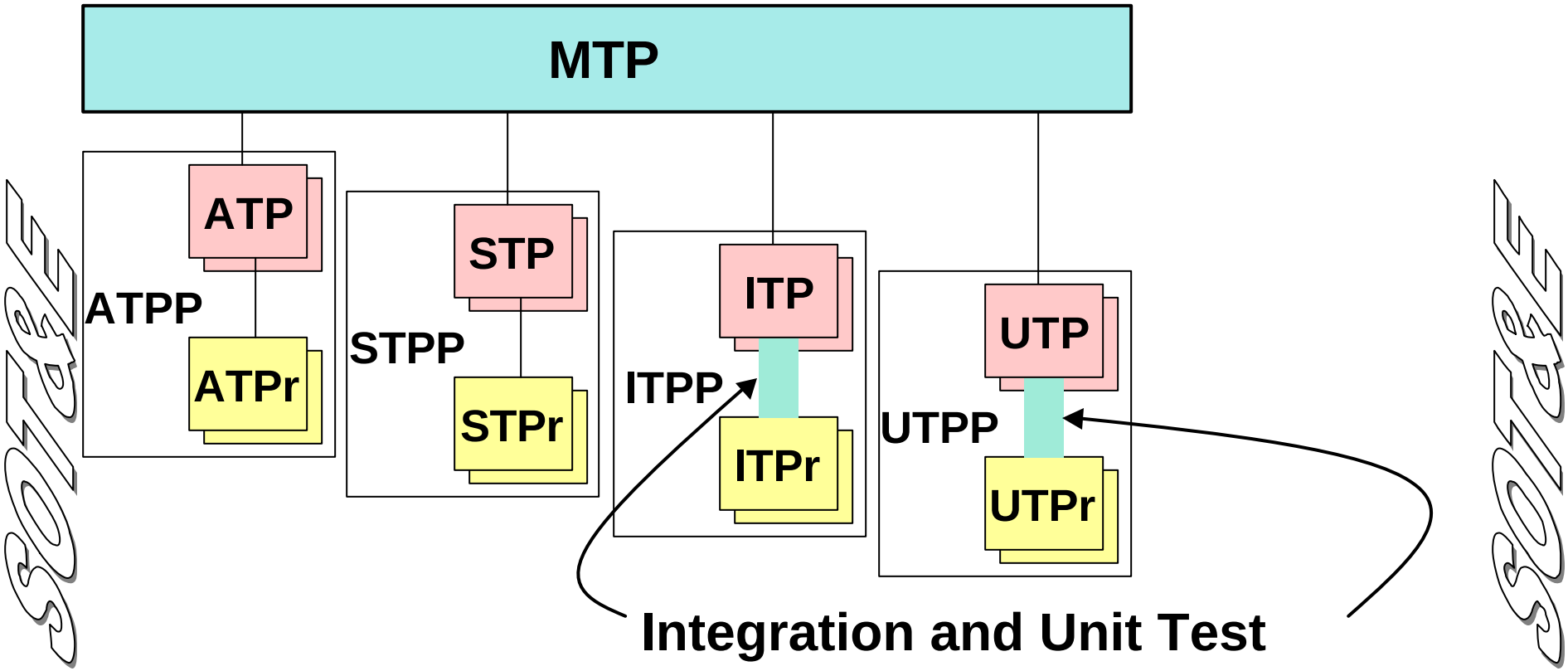
Test Planning

- **Test Plans and Test Procedures**
  - **Overall strategy defined in MTP**
  - **Test strategy for each level is defined in the level's test plan**

Acronyms for the next slide

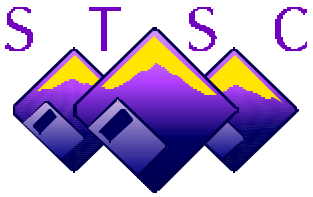
Test Level	Plan	Procedures	Plans and Procedures
Master	(MTP)		
Acceptance	(ATP)	(ATPr)	(ATPP)
System/Software	(STP)	(STPr)	(STPP)
Integration	(ITP)	(ITPr)	(ITPP)
Unit	(UTP)	(UTPr)	(UTPP)

Test Planning



Integration and Unit Test Plans and Test Procedures are often in the same document (i.e., ITPP or UTPP)

Test Case Design



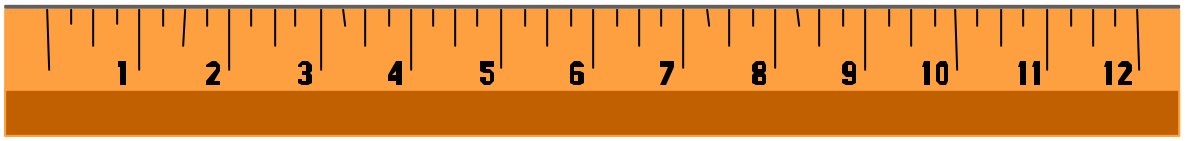
Area of Application	Test Case Design Techniques to Apply
Description of feature	Functional analysis
Specification with decision logic	Black-box path analysis
Unit test of new or revised code	Black-box and White-box path analysis
Numeric inputs	Boundary analysis
Inputs that are expected to produce the same type of output	Equivalence classes
Large populations of similar items	Statistical sampling
Large number of variables that affect the testing	Test factor analysis
Functions in specifications with decision logic	Cause-effect graphing
Frequency of use patterns	Operational profiling
Risk patterns	Risk assessment

Test Execution

- Test-Related Documentation
- Replication of Failures
- Results Evaluation
- Defect Buy-In from Developers
- False Test Results
- Localizing & Generalizing Defects
- Problem Reporting Issues
- Test Starting and Completion Criteria

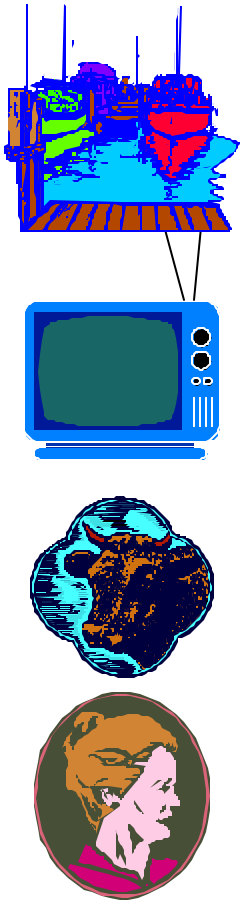
Test Management

- **Measures to monitor and track**
  - **Test Plan Development**
    - Actual Effort and Schedule vs Planned
  - **Test Procedure Development**
    - Actual number of test cases and test procedures developed vs Planned
    - Actual Effort and Schedule vs Planned
  - **Unit, Integration, and System Tests completed vs planned:**
    - And defects found and fixed by severity



Test Process Definition

- **A well-defined process includes:**
  - **P urpose**
  - **I nputs**
  - **E ntry Criteria**
  - **R oles**
  - **T asks**
  - **V erification**
  - **O utputs**
  - **e X it Criteria**
  - **M etrics**
  - **A uthority**

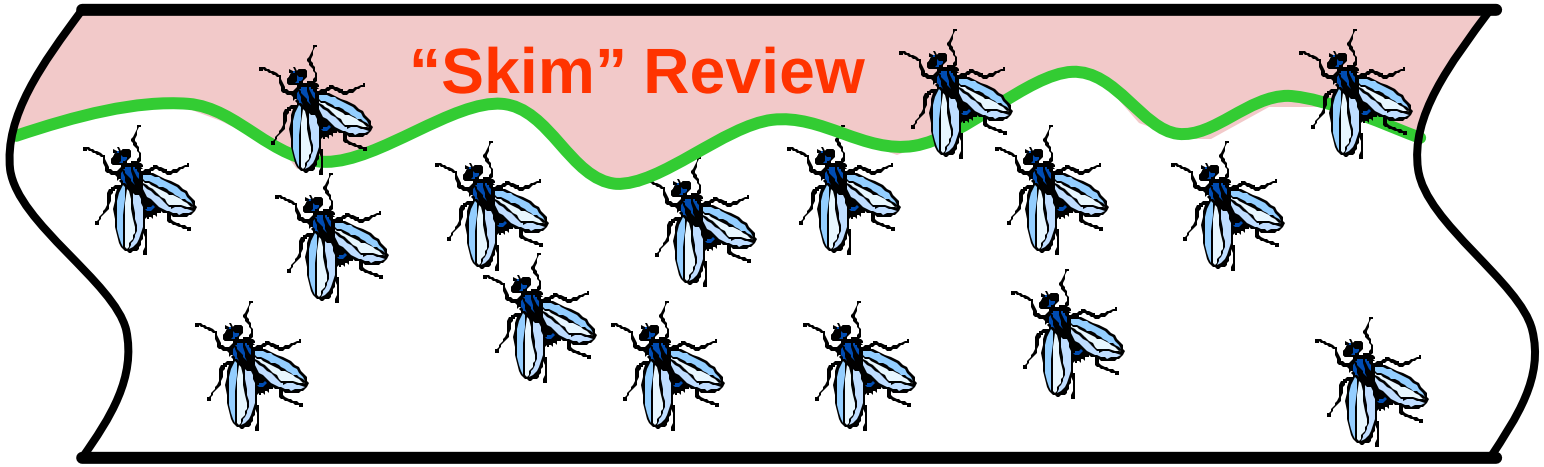


Should be:

- Be graphical with appropriate supporting textual detail
- Allow the process to be viewed in top-down (i.e., hierarchical like a WBS)
- Capture all desired process elements
- Have minimal redundancy
- Be easily understood
- Have the output available on-line
- Allow the output to be tailored easily for a project

[PERK98]

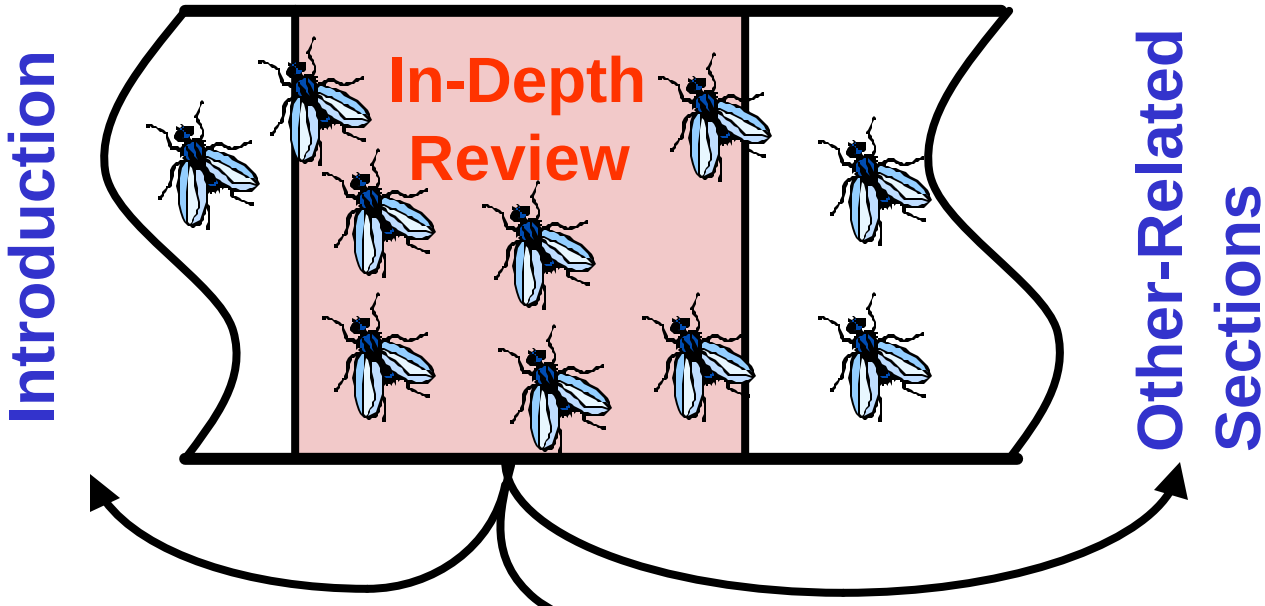
Test-Related Document Reviews



- **Results of a skim review are good as entrance criteria for an in-depth, Disciplined Document Review**
- **Skim review is most often used to overview**
- **Don't depend on ad-hoc, skim types of reviews to find defects**

Test-Related Document Reviews

Disciplined Document Reviews



Consider *sampling* if there isn't enough time to conduct in-depth review of entire document

Rules/Standards
Checklists
Source documents

Test Automation

- Test Case Generation
- Test Data Generation
- Complexity Measurement
- Debuggers
- Coverage Analysis
- Capture Replay Tools
- Comparators
- Load Testing Tools
- Test Managers
- Bug Trackers

Fundamental Software Best Practices

- Regression testing
- Developer unit-level testing
- Independent system-level testing
- Effective test case design techniques
- Risk analysis to drive testing
- Refined schedules that identify tasks / work products to be reviewed / tested prior to approval
- Software personnel trained
- Software test tools available and matured
- Effective configuration management practices
- Get involved early

Fundamental Software Best Practices

- **Software Program Manager's Network (SPMN)**
  - **Project Integrity**
    - Continuous risk manage
    - Metrics to manage
    - Earned value
    - Empirical cost & schedule estimation
    - People are the key resource
    - Defect tracking against quality targets
  - **Construction Integrity**
    - Life cycle config. management
    - System-based software design
    - Interfaces definition & control
    - Reuse risk and cost assessment
    - Requirements management / tracing
    - Data/database interoperability
    - Design twice, code once
  - **Product Stability, Integrity**
    - Require. & design inspection
    - Compile & smoke test frequently
    - Continuous testing process management

    [EVAN01]

Acronyms

- AF — Air Force
- AFOTEC — Air Force Operational Test and Evaluation Center
- AOA — Analysis of Alternatives
- ATP — Acceptance Test Plan
- ATPP — Acceptance Test Plans and Procedures
- ATPr — Acceptance Test Procedures
- ATPS — Automated Test Planning System
- COI — Critical Operational Issue
- CONOPS — Concept of Operations
- COTS — Commercial-off-the-Shelf
- CTP — Critical Technical Parameter
- DDR — Disciplined Document Review
- DoD — Department of Defense
- DT — Developmental Test
- DT&E — Developmental Test and Evaluation
- EOA — Early Operational Assessment
- FOC — Full Operational Capability
- FOT&E — Follow-On Test and Evaluation

Acronyms

- FRP — Full Rate Production
- HWCI — Hardware Configuration Item
- IOC — Initial Operational Capability
- IOT&E — Initial Operational Test and Evaluation
- IPT — Integrated Product Team
- ITP — Integration Test Plan
- ITPP — Integration Test Plans and Procedures
- ITPr — Integration Test Procedure
- KPP — Key Performance Parameter
- LFT&E — Live Fire Test and Evaluation
- LRIP — Low Rate Initial Production
- MCO — Marine Corp Order
- MNS — Mission Needs Statement
- MOE — Measures of Effectiveness
- MOP — Measures of Performance
- MOS — Measures of Suitability
- MTP — Master Test Plan
- NBCC — Nuclear, Biological, and Chemical Contamination

Acronyms

• NDI	Non-Developmental Item
• NDR	Need Determination Record
• OPP	Operational Performance Parameter
• ORD	Operational Requirements Document
• OT&E	Operational Test and Evaluation
• OTA	Operational Test Authority or Agency
• PDR	Preliminary Design Review
• PSA	Principle Staff Assistant
• RCM	Requirements Correlation Matrix
• RFP	Request for Proposal
• RVTM	Requirements Verification Traceability Matrix
• SAD	Software Architecture Description
• SAIC	Science Applications International Corporation
• SDD	Software Design Description
• SDP	Software Development Plan
• SECNAVINST	Secretary of the Navy Instruction
• SOT&E	Software-Oriented Test and Evaluation
• SOW	Statement of Work

Acronyms

- SPD — Software Product Description
- SPMN — Software Program Manager's Network
- SRD — Software Requirements Description
- SRS — System Requirements Specification
- STA — System Threat Assessment
- ST&E — Software Test and Evaluation
  - " — Security Test and Evaluation
- STC — Software Technology Conference
- STP — Software Test Plan
  - " — System Test Plan
- STPP — Software Test Plans and Procedures
- STPr — System Test Procedures
- STSC — Software Technology Support Center
- SWCI — Software Configuration Item
- T&E — Test and Evaluation
- TEMP — Test and Evaluation Master Plan
- TRR — Test Readiness Review
- UTP — Unit Test Plan

Acronyms

- UTPP Unit Test Plans and Procedures
- UTPr Unit Test Procedures
- V&V Verification and Validation
- VRTM Verification Requirements Traceability Matrix

References

[10 USC Sec139]	Title 10, U.S. Code, Section 139, Director, Operational Test and Evaluation, 29 Sept. 2000, Defense Acquisition Deskbook.
[AFI99-101]	AF Instruction 99-101, “Developmental Test and Evaluation,” 1 November 1996.
[AFI99-102]	AF Instruction 99-102, “Operational Test and Evaluation,” 1 July 1998.
[AFPD99-1]	Air Force Policy Directive 99-1, Test and Evaluation, “Test and Evaluation Process” 22 July 1993.
[AFOMan99-102]	AFOTEC Manual 99-102, “Operational Test Product Delivery Process Manual,” 20 January 2000.
[AFOPam99-102v2]	“AFOTEC Pamphlet 99-102, Volume 2 -- Software Support Life Cycle Process Evaluation Guide,” 1 Aug 1994.
[AFOPam99-102v4]	“AFOTEC Pamphlet 99-102, Volume 4 -- Software Usability Evaluation Guide,” 15 June 1994.
[AFOPam99-102v5]	“AFOTEC Pamphlet 99-102, Volume 5 -- Software Support Resources Evaluation Guide,” 28 Aug 1995.
[AFOPam99-102v6]	“AFOTEC Pamphlet 99-102, Volume 6 -- Software Maturity Assessment Guide,” 1 Mar 1996.

References

[AFOPam99-102v8]	"AFOTEC Pamphlet 99-102, Volume 8 -- Software Operational Assessment Guide," 15 June 1994.
[AR73-1]	Army Regulation, "Test and Evaluation Policy," 27 February 1995.
[ARPam73-7]	Dept. of the Army, "Software Test and Evaluation Guidelines," 25 July 1997.
[BEIZ90]	Beizer, Boris, Software Testing Techniques, 2nd edition, Van Nostrand Reinhold, 1990.
[CMMIC]	"CMMI-SE/SW, V1.0 Capability Maturity Model – Integrated for Systems Engineering/Software Engineering, Version 1.0, Continuous Representation," CMU/SEI-2000-TR-019, August 2000.
[COLL01]	Collard, Ross, "Systems Testing and Quality Assurance Techniques," Advanced Information Systems, Feb. 2001, Version 6.7.
[DAIC01]	Daich, Gregory T., "Software Test Process Improvement Workshop," STSC, Feb. 2001, Version 4.1.

References

[DoD5000.2-R]	DoD Regulation, “Mandatory Procedures for Major Defense Acquisition Programs (MDAPs) and Major Automated Information System (MAIS) Acquisition Programs,” Jan. 4, 2001.
[DoDD5000.1]	DoD Directive, “The Defense Acquisition System,” 23 Oct. 2000.
[DoDI5000.2]	DoD Instruction, “Operation of the Defense Acquisition System,” Jan. 4, 2001.
[DSMCGlossary]	Defense Systems Management College, Glossary; Defense Acquisition Acronyms and Terms, Ninth Edition, November 1998.
[EVAN01]	Evans, Michael W., “SPMN Director Identifies 16 Critical Software Practices,” CrossTalk, STSC, March 2001, Vol. 14 No.3.
[IEEE610.12]	IEEE Std 610.12-1990, “IEEE Standard Glossary of Software Engineering Terminology,” 15 Feb. 1991.

References

[J-STD-016]	EIA/IEEE J-STD-106-1995, Interim Standard, “Standard for Information Technology Software Life Cycle Processes Software Development Acquirer-Supplier Agreement,” 30 September 1995.
[MCO3960.2B]	MCO 3960.2B -- CIC040, “Marine Corps Operational Test and Evaluation Activity,” 24 October 1994.
[PERK98]	Perkins, Timothy, “Process Definition Workshop,” STSC, Oct. 5-6, 1998.
[PN99-502]	“Eagle Look Test and Evaluation of Software-Intensive Systems,” PN 99-502, Air Force Inspection Agency, 19 April 1999.
[SECNAVI5000.2B]	“Implementation of Mandatory Procedures for Major and Non-Major Defense Acquisition Programs and Major and Non-Major Information Technology Acquisition Programs,” 6 December 1996.
[TST101]	“TST-101 Introduction to Acquisition Workforce Test & Evaluation (IAWTE)”, DAU, March 2000.
[TST202]	“TST 202 Intermediate Test and Evaluation,” DAU, 4-12 Apr. 2000.